Was drin ist für dich: Algorithmen besser verstehen und lernen, was sie gefährlich macht.

Kennst du diese Momente, in denen man sich fragt: Leben wir eigentlich schon in der Zukunft? 1965 schrieb Philip K. Dick die später mit Tom Cruise in der Hauptrolle verfilmte Geschichte Minority Report. Die darin beschriebene Technologie, in der Verbrechen vorausgesehen werden können, bevor sie geschehen, war damals noch reine Science-Fiction.

Mittlerweile gibt es diverse Algorithmen, die darauf spezialisiert sind, Verbrechen vorauszuahnen. Aber wie genau arbeitet ein solcher Algorithmus? Und wollen wir uns wirklich von derartigen Technologien abhängig machen? In diesen Blinks geben wir dir Antworten auf diese Fragen und zeigen dir, warum wir eine Zukunft mit willkürlich herrschenden Computern verhindern sollten.

In den Blinks wirst du außerdem herausfinden,

wieso Algorithmus nicht gleich Algorithmus ist,
warum eine KI prinzipiell nichts ist als ein Computer, der ein bisschen herumprobiert, und
was wir tun können, um die KI-Dystopie zu verhindern.

Algorithmen sind die Zukunft – deshalb müssen wir uns mit ihrer Funktionsweise beschäftigen.

Wann hast du dir zuletzt Gedanken darüber gemacht, wie unabhängig du in deinem Alltag von Algorithmen bist? Wenn Netflix dir einen Film empfiehlt oder Amazon dir ein Produkt ans Herz legt, dann liegt dem ein Algorithmus zugrunde. Auch wenn du nach dem aktuell schnellsten Weg von Köln nach Hamburg suchst oder dir auf der Website von Zalando verschiedene Oberteile geordnet nach ihrem Preis anzeigen lässt, sind Algorithmen im Spiel.

Diese Systeme scheinen – zumindest auf den ersten Blick – recht unspektakulär. Doch im Verdeckten agieren weit weniger harmlose Algorithmen. Bei einigen Banken zum Beispiel wird dir ein Kredit nur

VISIT…

LANZAROTE Caliente.com

dann gewährt, wenn deine Kundendaten für den Algorithmus passen. Ein freundliches Gespräch mit deiner Sachbearbeiterin bringt dich hier also nicht mehr weiter, denn über deine Kreditwürdigkeit entscheidet kein Mensch. Manche Unternehmen nutzen mittlerweile Algorithmen, um geeignete Bewerber zu finden. Und es gibt sogar Algorithmen, die berechnen, wie wahrscheinlich Straftäter rückfällig werden.

Letztere gehören zu den sogenannten richtenden Algorithmen – sie sagen menschliches Handeln voraus und treffen anschließend Entscheidungen über uns. Vor solchen richtenden Algorithmen müssen wir uns in Acht nehmen. Ebenso vor dichtenden Algorithmen – was ja erst mal ein bisschen nach Roboter-Goethe klingt. Aber diese Algorithmen haben es tatsächlich in sich, denn sie sollen unsere Handlungen imitieren und optimieren – was bedeutet, dass sie uns auch potenziell ersetzen können. Ganz konkret lässt sich das an Übersetzungsalgorithmen zeigen. Sie sind heute schon so gut, dass sie in naher Zukunft Übersetzer und Dolmetscher ersetzen könnten.

Richtende wie dichtende Algorithmen gibt es immer mehr. Sie agieren hinter den Kulissen, sie treffen dort ihre Urteile – oft ohne dass wir uns darüber bewusst sind. Die alles entscheidende Frage dabei lautet natürlich: Wollen wir so leben? Und falls nicht: Was können wir tun, damit wir künftig nicht noch mehr Kontrolle und Aufgaben an Algorithmen abgeben?

Klar ist: Algorithmen gehört die Zukunft. Um die Kontrolle nicht zu verlieren, müssen wir die Technik und ihre Auswirkungen auf die Gesellschaft besser verstehen. Dazu hilft ein Blick in den „Maschinenraum“. Hier können wir uns genauer ansehen, wie Algorithmen aufgebaut sind und wie sie entscheiden. Du wirst erfahren, worin sie gut sind, aber auch, wieso sie bei Weitem nicht so objektiv und erst recht nicht so unfehlbar sind, wie du vielleicht glaubst.

Klassische Algorithmen lösen mathematische Probleme auf möglichst elegante Weise.

Im vorigen Blink hast du schon einige Arten von Algorithmen kennengelernt. Grundsätzlich lassen sie sich in zwei Kategorien einteilen: klassisch und selbstlernend.

Klassische Algorithmen funktionieren nach den Vorgaben eines Programmierers. Du kannst sie dir als eine Art Bauanleitung oder Fahrplan vorstellen. Ein klassischer Algorithmus löst dabei immer ein mathematisches Problem. Ein solches kann beispielsweise darin bestehen, den kürzesten Weg von A nach B zu finden oder etwas nach verschiedenen Aspekten zu sortieren.

Die wichtigste Gestaltungsleistung beim klassischen Algorithmus ist die Modellierung. Modellierung meint, ein mathematisches Problem so zu vereinfachen, dass der Algorithmus möglichst leicht und schnell eine passende Lösung finden kann.

Probleme können gut oder schlecht modelliert sein. Edsger Dijkstra, ein niederländischer Programmierer, fand Mitte der Fünfzigerjahre beispielsweise eine sehr elegante Modellierung für das Problem des kürzesten Weges. Dazu stellte er sich eine Straßenkarte als Netzwerk vor. Kreuzungen wurden darauf zu Knotenpunkten eines Netzes, die durch Straßen verbunden sind, aber auch Ortschaften entlang einer Strecke machte er zu Knoten. Nach dieser Modellierung, die sich Dijkstra vor knapp 70 Jahren angeblich in nur zwanzig Minuten ausdachte, funktionieren Navigationssysteme noch heute.

Das zweite entscheidende Moment beim Codieren von Algorithmen besteht in der Operationalisierung. Dabei geht es darum, einen unmathematischen Bereich messbar zu machen. Im Falle von Dijkstras Kürzester-Weg-Algorithmus war das die Länge der Strecke in Kilometern. Auch ein einfaches Konzept wie dieses muss für den Algorithmus messbar gemacht werden – schließlich ist nicht zwingend die tatsächliche Kilometerzahl ausschlaggebend.

Möglicherweise interessiert sich ein Nutzer mehr für die erwartete Fahrzeit.

Etwas zu operationalisieren ist jedoch nicht immer so leicht. Denn Geschmack, Kreditwürdigkeit oder sogar Liebe zu vermessen ist eine deutlich größere Herausforderung.

Aus Operationalisierung und Modellierung ergibt sich das sogenannte OMA-Prinzip. Die Ergebnisse, die ein Algorithmus ausspuckt, können immer nur im Rahmen der Modellierung und entsprechender Operationalisierungen sinnvoll interpretiert werden. Der Kürzeste-Weg-Algorithmus zum Beispiel wird bei einer Bahnreise fehlerhafte Ergebnisse liefern. Bei einer Reise mit der Bahn müssen nämlich auch andere Werte wie etwa Umstiegszeiten mit in die Modellierung einbezogen werden.

Klassische Algorithmen haben eine „Superkraft“: Ihr Funktionieren ist nicht abhängig von großen Datenmengen. Und hier wird der Unterschied zum selbstlernenden Algorithmus deutlich.

Selbstlernende Algorithmen probieren herum – was mal besser, mal schlechter klappt.

Nicht alle Probleme lassen sich mit einem klassischen Algorithmus lösen. Klassische Algorithmen leisten das, was Menschen vorher programmiert haben – nicht mehr und nicht weniger. Die Vorhersage menschlichen Verhaltens wurde erst durch selbstlernende Algorithmen denkbar.

Ein besonders weit verbreiteter selbstlernender Algorithmus ist der Empfehlungsalgorithmus. Ob Netflix, Amazon oder Zalando: Wenn dir online jemand etwas andrehen will, benutzt er sicher einen solchen Algorithmus. Um Empfehlungsalgorithmen zu verstehen, musst du dir klarmachen, dass sie eigentlich nur durch Herumprobieren zu ihren Ergebnissen kommen.

Der Fachbegriff für algorithmisches Herumprobieren lautet übrigens Heuristik. Etwas technischer ausgedrückt, ist eine Heuristik eine

Strategie, mit der man für ein Problem eine gangbare Lösung findet. Die am weitesten verbreitete Methode ist dabei, schon vorhandene Daten nach Mustern zu durchsuchen, um diese auf die Zukunft zu projizieren. Frei nach dem Motto: „Wer den Film Herr der Ringe mochte, wird auch die Serie Game of Thrones mögen.“

Selbstlernende Algorithmen sind klassischen Algorithmen insofern überlegen, dass sie unfassbar große Datenmengen auf korrelierende Aspekte durchsuchen können. Das Lernen funktioniert so, dass der Algorithmus nach jedem Datenpaket Feedback zu seinen bisherigen Ergebnissen erhält. Ganz ähnlich, wie auch Kinder lernen. Aus diesem Grund spricht man bei dieser Art von Algorithmus auch von künstlicher Intelligenz (KI). Klassische Algorithmen kommen hingegen schon fertig auf die Welt und lernen nichts mehr dazu.

Doch – und das ist der entscheidende Punkt – es sind nicht zwangsläufig die besten Lösungen, auf die eine KI kommt. Unter Umständen findet sie nur wenig aussagekräftige oder gar sinnlose Muster. Gerade im kommerziellen Bereich reicht jedoch oft auch schon eine annähernd richtige Lösung. Bei Filmempfehlungen sind Korrelationen wie im Falle von Game of Thrones und Herr der Ringe schon einigermaßen gut. Und je mehr Daten er bekommt, desto besser wird der selbstlernende Algorithmus. Eine schwache Korrelation beispielsweise kann durch große Datenmengen ausgeglichen werden.

Problematisch wird es, wenn man eine solche KI für mehr als Film- oder Kaufempfehlungen einsetzt – im naiven Glauben, sie könne objektive Lösungen bieten. Bevor wir uns derartige Fälle anschauen, zeigen wir dir im nächsten Blink erst einmal, worin selbstlernende Algorithmen richtig gut sind.

Selbstlernende Algorithmen übertreffen Menschen schon in vielen Bereichen.

Hast du in letzter Zeit einen Zoo besucht und warst dir vor dem Gehege stehend unsicher, was du da gerade eigentlich siehst? Hast du dich auf dem Weg zur Informationstafel gefragt, ob du da gerade einen Seelöwen, einen Seebär oder einen Seehund vor dir hast? Schwierig wird es immer dann, wenn Experten – in diesem Fall Tierforscher – einer Sache viele kleine Unterkategorien mit spezifischen Namen zuordnen. Im Alltag braucht man solche Unterscheidungen nämlich nur selten. Wenn du Angst vor einer am Fußboden herumkriechenden Spinne bekommst, wird es dir schließlich egal sein, ob es sich um eine Feld-, Mauer- oder Waldwinkelspinne handelt.

Die Beispiele zeigen: Es ist gar nicht so leicht, alle Dinge, die wir sehen, der exakt richtigen Kategorie zuzuordnen. Genau darin sind selbstlernende Algorithmen aber besonders gut. Und sie werden immer besser. Zwischen 2010 und 2017 sank die Fehlerrate in der Bilderkennung durch selbstlernende Algorithmen von achtundzwanzig auf zwei Prozent. Menschen sind da weit weniger sicher, sie ordnen durchschnittlich fünf Prozent der Bilder falsch zu.

Die Möglichkeiten, die sich mit dem Einsatz von Bilderkennungssoftware erschließen, sind vielfältig. Ein zentrales Anwendungsgebiet sind selbstfahrende Autos. Eine KI, die einen Wagen steuert, muss sicher erkennen können, ob sich das Kind am Straßenrand in Richtung Straße bewegen könnte – zum Beispiel, weil es Ball spielt – oder ob dahingehend keine Gefahr besteht.

Bilderkennungsalgorithmen sind bereits weit fortgeschritten. Bei der Erkennung von Hautkrebs kommen sie teilweise schon zum Einsatz, denn auch darin sind sie sehr gut. So zeigte eine Studie aus dem Jahr 2018, dass ein selbstlernender Algorithmus bösartige Tumoren mithilfe von Bildern besser erkennen konnte als Dermatologen.

Dafür wurde der Algorithmus trainiert, anhand von Bildern gutartige Melanome von bösartigen – also Hautkrebs – zu unterscheiden. Im Wesentlichen wurden die Bilder dabei nach Krebsrisiko sortiert. Anschließend wurden auch 58 Hautärzte gebeten, die Bilder zu

beurteilen. Das Ergebnis: Dermatologen erkannten zwar die bösartigen Melanome, doch sie ordneten häufiger als die Software gutartige dem Krebs zu.

Selbstlernende Algorithmen können nur dann gut funktionieren, wenn man ihnen einen klar definiertes Qualitätsmaß an die Hand gibt und sie mit großen Datenmengen füttert. Was aber, wenn diese Dinge fehlen?

Diskriminierung ist eines der größten Probleme in der Anwendung von selbstlernenden Algorithmen.

Sichere Fahrten mit autonomen Fahrzeugen, eine exzellente Krebserkennungsrate: selbstlernende Algorithmen können uns in vielen Bereichen sehr helfen.

Aber es gibt auch Anwendungsbereiche, in denen selbstlernende Algorithmen Probleme bereiten – etwa, weil sie diskriminieren oder vorhandene Diskriminierungen noch weiter verstärken. Ein gutes Beispiel dafür ist die Software COMPAS. Sie wird in den USA von Gerichten eingesetzt und soll bewerten, mit welcher Wahrscheinlichkeit Straftäter rückfällig werden. Der Algorithmus vergibt dafür einen Risikoscore – ähnlich wie bei den bereits erwähnten Krebsbildern.

Aber was bedeutet das eigentlich, rückfällig werden? Straftat ist schließlich nicht gleich Straftat. Manche Arten von Verbrechen werden häufiger aufgedeckt als andere. So werden in den USA sogenannte White Collar Crimes wie Steuerhinterziehung weitaus seltener aufgeklärt als Drogendelikte. Da gewisse Verbrechen auffälliger sind, wirkt sich das auch auf den Risikoscore aus: Manche Menschen werden nur aufgrund der Art ihrer Vergehen schlechter bewertet als andere.

Die Ironie dabei ist, dass man sich von Algorithmen ursprünglich eine unvoreingenommene Bewertung erhofft hatte. Die Annahme war: Weil Menschen subjektiv über andere Menschen urteilen, verwendet man einen objektiven Algorithmus.

Doch zu glauben, Algorithmen würden objektiv entscheiden, ist naiv. Nehmen wir einmal Amazons selbstlernenden Bewerbungsalgorithmus. Dieser wurde so programmiert, dass bestimmte Merkmale wie Geschlecht, Alter und Nationalität nicht berücksichtigt wurden, um möglichst unvoreingenommen zu urteilen. Das Problem: Der Algorithmus fand trotzdem Merkmale, die mit dem Geschlecht korrelierten. So bewertete die Software Abschlüsse von Colleges, die ausschließlich Frauen zulassen, schlechter. Dazu kann es kommen, wenn ein selbstlernender Algorithmus beispielsweise feststellt, dass Menschen mit dem Abschluss eines solchen Colleges zuvor mit einer geringeren Erfolgsquote eingestellt wurden. Stichwort: falsche Korrelation.

Diskriminierung geschieht mitunter aber auch ganz einfach, weil zu wenig Daten von bestimmten Menschengruppen vorhanden sind. Wenn in Zukunft Spracherkennungssoftware immer entscheidender wird, kann das dazu führen, dass Menschen mit starkem Akzent sehr schlecht mit der Software zurechtkommen.

Der Mensch entscheidet, wie die Ergebnisse von Algorithmen interpretiert werden.

Hast du dir beim Lesen des vorigen Blinks vorgestellt, wie sich die falsche Zuordnung durch eine KI anfühlen würde? Möglicherweise hast du dich auch über eine gewisse Arroganz geärgert, mit der unausgereifte Algorithmen über deine Kreditwürdigkeit oder deine Qualifizierung entscheiden, als seien sie Götter, die stets objektiv richtige Entscheidungen fällen könnten.

Vor allen Dingen ist es grob fahrlässig, die Erfolge von Algorithmen im kommerziellen Bereich einfach so auf andere Problemfelder wie das Handeln von Menschen zu übertragen. Keine KI der Welt ist nämlich imstande, absolut gültige Regeln für menschliches Verhalten zu finden, geschweige denn, es zweifelsfrei vorherzusagen.

Eine Heuristik, die Daten auf Korrelationen überprüft und Menschen so Gruppen zuordnet, sagt immer nur Wahrscheinlichkeiten voraus.

Das ist jedoch zu wenig, wenn es um wichtige gesellschaftliche Ressourcen geht und man keine Möglichkeit hat, zu widersprechen. Wie wir uns auf einer individuellen Ebene verhalten, hängt darüber hinaus auch nicht nur von unserem Charakter ab. Kontext ist alles: Wir sind immer beeinflusst von der Situation, in der wir uns gerade befinden.

Algorithmen sind nur Werkzeuge. Sie haben keine Vorurteile, Menschen aber schon. Und es sind gerade diese Menschen, die die von Algorithmen gelieferten Ergebnisse interpretieren und darüber entscheiden, wie man mit den Fehlern von Algorithmen umgeht.

Je nachdem, wo man eine Trennlinie zieht, kommen mehr oder weniger unbescholtene Menschen ins Gefängnis. Dies ist jedoch keine technische Frage, hier geht es um Moral. Kein Programmierer oder Datenforscher kann bestimmen, an welcher Stelle die Grenze zwischen Schuld und Unschuld gezogen wird. Bei Algorithmen, die über uns entscheiden, sind wir alle gefragt. Die Ethik findet erst über uns in den Rechner.

Wie genau wir wirklich mitbestimmen können und welche Strukturen die Politik dafür schaffen müsste, schauen wir uns im nächsten und letzten Blink an.

Wir müssen die Kontrolle über die KI behalten.

Es bleibt somit nur noch eine große Frage: Was können wir tun, um Missbrauch durch Algorithmen zu verhindern? Die Antwort darauf scheint relativ klar: Es sollte in unserer Hand liegen, zu steuern, wo Algorithmen sinnvoll sind – und wo nicht. Wir brauchen feste Regeln, die uns als Bürger vor diskriminierenden KI-Systemen schützen.

Eine staatliche Regulierung ist vor allem da vonnöten, wo selbstlernende Algorithmen über Menschen richten. Sobald ein Algorithmus über unseren Zugang zu wichtigen gesellschaftlichen oder ökonomischen Ressourcen entscheidet, über Kredite und Einladungen zu Bewerbungsgesprächen, müssen wir uns

einmischen. Hierzu fehlen bisher jedoch noch unabhängige Prüfinstitutionen.

Natürlich heißt das nicht, dass von nun an jeder Algorithmus überprüft werden soll. Sinnvoller ist eine sogenannte Risikomatrix, mit der relativ harmlose Algorithmen vorab ausgesiebt werden. Mit einer solchen Matrix ließe sich sehr leicht verbildlichen, bei welchen Systemen wir als Gesellschaft eingreifen müssen, um einem Missbrauch vorwegzugreifen.

In dieser Matrix wird zwischen fünf Klassen unterschieden – sie reichen von 0 bis 4. Systeme, die einfach nur Produkte empfehlen, gehören der Klasse 0 an. In einer höheren Klasse befinden sich dann Algorithmen wie der sogenannte China Citizen Score, bei dem das Verhalten chinesischer Bürger gemessen und anschließend mit einer einzigen Zahl bewertet wird. In der vierten und höchsten Risikoklasse befinden sich algorithmische Entscheidungssysteme, deren Schadenspotenzial als extrem hoch angesehen wird oder die rechtlich nicht durchsetzbar sind; sie sollten am besten gar nicht existieren.

Auch beim Datenschutz muss sich etwas tun. Unternehmen oder Regierungen, die über große Datenmengen verfügen, gewinnen mit diesen Einblick in unser Leben. Mit großen Datenmengen lassen sich Persönlichkeitsprofile erstellen, die uns in Zukunft den Zugang zu bestimmten Ressourcen versperren könnten.

Die einfachste Art und Weise, wie wir all das erreichen, ist mit einer neuen Generation von Entwicklerinnen und Entwicklern, die sich dieser Probleme bewusst sind. Im Modellstudiengang Sozioinformatik, der an der TU Kaiserslautern unterrichtet wird, lernen die Studierenden neben dem Programmieren, dass Algorithmen immer Teil eines großen Ganzen sind. Es sind schließlich die Algorithmen, die sich der Gesellschaft anpassen müssen, nicht umgekehrt.

Zusammenfassung

Die Kernaussage dieser Blinks ist:

Zwar sind selbstlernende Algorithmen in einigen Bereichen besser als Menschen, wie beispielsweise in der Bilderkennung. Doch sie sind nicht für alle Fälle gleich gut geeignet. Wir müssen aufpassen, dass sie nicht dort über uns richten, wo wir es gar nicht wollen. Gefährlich wird es, wenn wir wegen eines Algorithmus nicht eingestellt werden oder kein Bankkonto eröffnen können: Diskriminierung ist eines der größten Probleme in der Anwendung von selbstlernenden Algorithmen. Um sie zu verhindern, müssen wir klare Regeln aufstellen, wo welche KI eingesetzt werden darf. Nur so können wir zukünftig die Kontrolle über künstliche Intelligenz behalten.